Views among Economists: Professional Consensus or Point-Counterpoint?*

by

Roger Gordon and Gordon B. Dahl
(UCSD)

To what degree do economists disagree about key economic questions? Our self-image as a profession would be that our views on economic questions are based on the accumulated academic evidence, both theoretical and empirical. As a literature first develops, the evidence will still be ambiguous, with both alternative theories and contradictory empirical estimates, easily leading to disagreements. But as evidence accumulates, a professional consensus should emerge, perhaps more quickly among those working in the relevant field.

An alternative perception, reflecting the traditional fresh-water/salt-water divide in macroeconomics, is that economists coalesce into different camps, to a degree reflecting a liberal/conservative divide, with one group focusing on evidence that government intervention is almost always too costly ex post to be justified and another that market failures are all too frequent and can be alleviated by well-designed policy interventions. Compounding this divide is differences in the degree of importance given to distributional vs. efficiency implications of alternative policies.

These camps have traditionally been associated with particular universities, with Chicago, Rochester, UCLA, and Stanford arguably reflecting the conservative camp, and MIT,

---

* Contact information: Roger Gordon (contact author), Department of Economics, UCSD, 9500 Gilman Dr. #0508, La Jolla, CA 92093-0508; rogordon@ucsd.edu. Gordon B. Dahl, Department of Economics, UCSD, 9500 Gilman Dr. #0508, La Jolla, CA 92093-0508; gdahl@ucsd.edu. We would like to thank Anil Kashyap for inspiring us to write this paper, and both Anil Kashyap and Brian Barry for comments on an earlier draft.

Harvard, and Berkeley reflecting the more liberal camp. Such a division could also show up among Ph.D.'s from different schools, who each received training from a different set of faculty.

There may also be generational divides among economists, with a changing consensus reflecting not so much each of us changing our views based on the accumulating evidence, but instead those with out-dated views gradually retiring and being replaced by younger economists whose views rely on more recent evidence.

This perception of different camps among economists is reinforced by the use of economic spokespeople for Presidential candidates and sitting Presidents. The job of such a spokesperson is to defend the politician's positions in public, regardless of the economic advice the spokesperson is giving internally, assuring a perception of politicized views. This perception is reinforced by the media's use of economists with polarized views in their point-counterpoint debates.

These spokespeople are also under pressure to be "one-armed" economists, avoiding the many qualifications coming out of the academic literature. Even non-spokespeople can be subject to an "expert bias", a bias towards more certainty than is justified, with this bias perhaps stronger for men than women.

A hybrid view would be that economic priors (conservative vs. liberal) matter when the empirical evidence is weak, but not when the empirical evidence is strong. The degree of polarization would then depend (negatively) on the depth of the associated economic literature.

To provide evidence on the degree to which there is a consensus vs. polarized camps of economists, we make use of the responses to a series of questions posed to a distinguished panel of economists put together by the Chicago Booth School of Business, using responses up through October 30, 2012. By design, the top seven Departments are equally represented on the Panel.

The Panel includes economists from different cohorts, women as well as men, and members of different political persuasions. Each week since late September, 2011 this Panel has been asked to respond to a statement expressing a particular view on an economic question of current interest. Panel members could “strongly agree”, “agree”, be “uncertain”, “disagree”, or “strongly disagree” with the statement, or alternatively decline to express an opinion or simply not answer. Members also indicated their confidence in their response, on a scale from one to ten.

The aim of this paper is to examine these responses, to see to what degree opinions differ within the Panel. To what degree is there more consensus among those from a particular school, or those from a particular cohort or a particular gender? When opinions differ, does this division broadly correspond to a liberal/conservative divide? Is there less disagreement on topics with a large academic literature?

Our null hypothesis is that economists have homogeneous views, informed by whatever academic evidence exists. The evidence may be ambiguous, but economists would then agree on this ambiguity. Under our alternative hypothesis, we expect to see diverse views, perhaps linked to a liberal/conservative divide, and perhaps with less diverse views *within* Departments or cohorts and when the academic literature is large.

Based on our analysis, we conclude that there is close to full consensus among these Panel members when the past economic literature on the question is large. When past evidence is less extensive, differences in opinions do show up. But there is no tendency for those with the same gender, from the same cohort, from the same Department, or with Ph.D.’s from the same school, to have similar views. There are certainly some idiosyncratic views expressed, but we found no evidence of different camps. We did find, though, that women and younger economists are less willing to express an opinion, while those with experience working in Washington have fewer such inhibitions. On net, the main finding is of a broad consensus on these many different economic issues.

The description of the data and of our empirical methods is described in the next section, along with the resulting estimates. Section II then provides a brief summary.

I. Empirical Methods

The first question we face in the analysis is how to define differences in views among our panel members. Consider for example the responses to the question asking whether the Fed's new policies in 2011 will increase GDP growth by at least 1% in 2012. The responses were 41% "uncertain", 37% "disagree", and 17% "strongly disagree". Initially, we assume that these responses imply consensus, since no respondents "agree" even though respondents differed in their extent of disagreement with the statement.

By this definition, there is an extraordinarily high level of consensus among our Panel members.[1] Only 6% of the responses in our data violate this definition of consensus, responding "disagree" when the consensus is "agree", or conversely.[2] Similarly, for 32 out of the 80 questions, there are no such disagreements.

Disagreements are not confined to questions where one might expect polarized views. The greatest disagreement, for example, arose with the following statement from May 23, 2012, where opinions were almost equally divided between "agree", "disagree", and "uncertain" (or

[1] Throughout our analysis, we group together "agree" with "strongly agree", and "disagree" with "strongly disagree". We also group together "uncertain" with "no opinion", finding these two responses have very similar correlations with other observables. In contrast, "did not answer", while correlated with gender and age, had no significant correlation with question characteristics, suggesting nonresponses when respondents were simply too busy rather than when they were unsure of the best response.
[2] If we instead define consensus to be identical answers among the three categories, we still find only minor disagreements: only 30% of the responses deviate from the modal response for a question.

"no opinion"): "New technology for fracking natural gas, by lowering energy costs in the United States, will make US industrial firms more cost competitive and thus significantly stimulate the growth of US merchandise exports." Yet there is no obvious partisan division on this issue. What we'll argue later is that such disagreements are much more common when the academic literature on an issue is small, as is the case for "fracking".

In spite of these observed differences in responses, though, views among our Panel might still be homogeneous. This could occur if each individual responds randomly, reflecting ambiguities in the evidence and disagreement with some elements of a statement but agreement with other parts. To allow for such random responses, we now assume under the null hypothesis that the *probabilities* of any of the three possible answers to a given question are the same for all respondents.

To test whether all individuals have the same probabilities, we first tabulated the number of times each individual agreed with the consensus, was uncertain, or disagreed with the consensus.[3] We modeled these counts as coming from a multinomial distribution, with probabilities differing by question, and conducted a chi-square test of homogeneity to judge whether the distribution of responses was consistent with individuals having a common set of response probabilities.[4]

The resulting test strongly rejects homogeneous responses, with a p-value less than 0.001. Deviations, though, take many forms: some individuals are much more likely to disagree with the consensus, some are more likely to agree, and others are more likely to be uncertain.

To what degree are these deviations linked to observable characteristics of the Panel

[3] These tabulations are available in the Online Appendix.
[4] In particular, we simulated the distribution of the chi-squared statistic under the null hypothesis, bootstrapping the estimated probability of each response for each question, and then measured the probability of seeing a chi-squared statistic as extreme as that implied by the actual responses.

members? (Under the null hypothesis, individual responses to each question would be entirely random, given the overall probabilities in the Panel as a whole.) To provide some evidence, we estimated a multinomial logit model forecasting the probability of each response as a function of the characteristics of each respondent. The characteristics we included were:[5] a dummy for female, vintage of Ph.D.,[6] a dummy indicating experience working in Washington, and a dummy indicating expertise about the specific topic.[7]

Results are reported in Table 1. To aid in interpretation, the table provides marginal effects evaluated at the means of the other covariates. In the first specification, we find that women are 8 percentage points more likely to be uncertain compared to men. This is a sizeable effect given that one-fourth of individuals answer "uncertain" on average. None of the other coefficients are statistically significant. Using a likelihood ratio test for the joint significance of all the personal characteristics other than effects of gender on being "uncertain", we find a p-value of 0.73, showing no systematic differences in views other than women being less likely to express an opinion.

This first specification has not controlled for question characteristics, which could be important since some questions evoke more consensus than others. We therefore added the following additional controls to capture systematic variation by question in the fraction giving each response: a dummy indicating that the response follows immediately from undergraduate price theory, a dummy indicating a large academic literature on the topic, a dummy indicating at

[5] In an additional specification, we included dummy variables indicating the university where the respondent got his or her Ph.D. as well as the university where they are currently employed. The resulting coefficients were not individually or jointly significant, and are omitted from the reported specification.
[6] Cohorts are defined to equal: zero to fifteen years since Ph.D., sixteen to thirty years since Ph.D., and greater than thirty years since Ph.D.
[7] Here, we assigned each question to a specific field or fields. An individual is an expert if there is any overlap among the fields covered by the question and the individual's primary field of membership at NBER. The specific assignments are available in the Online Appendix. For the few panel members who were not affiliated with NBER, we "assigned" them to the obvious programs.

Table 1. Multinomial Logit Estimates of the Probability of "Uncertain" and "Reject Consensus" Relative to "Agree with Consensus".

Variable	Marginal Effect	t-test	Marginal Effect	t-test	Marginal Effect	t-test
Uncertain (average rate = 0.245)						
Female	.083	3.89	.084	3.93	.072	3.81
Washington	.013	0.65	.016	0.81	.019	1.04
Vintage: older	-.016	-0.77	-.013	-0.63	-.009	-0.45
mid-career	-.010	-0.51	-.010	-0.52	-.003	-0.16
Expert	-.022	-1.21	-.004	-0.20	-.034	-2.04
Lit: int. micro			-.247	-11.38		
large			-.234	-10.84		
medium			-.048	-2.05		
Perfect markets?			-.030	-1.44		
Redistribution?			-.045	-1.89		
Avg. confidence					-.017	-2.38
Reject consensus (average rate = 0.060)						
Female	-.001	-0.06	.001	0.07	.002	0.23
Washington	.005	0.42	.006	0.58	.004	0.39
Vintage: older	.017	1.39	.015	1.33	.015	1.29
mid-career	.005	0.45	.004	0.38	-.000	-0.01
Expert	.010	0.98	.012	1.20	.007	0.72
Lit: int. micro			-.083	-6.82		
large			-.066	-4.95		
medium			-.034	-2.43		
Perfect markets?			.023	1.79		
Redistribution?			.008	0.62		
Avg. confidence					.010	2.34
Question dummies?	No		No		Yes	

Notes: 2,834 observations in each specification. The omitted category is a young male, without experience working in Washington, without any expertise on the topic, facing a question with no past academic literature. Marginal effects are evaluated at the means of the other variables.

least a few academic publications on the topic, an indicator for whether responses could reflect distributional concerns and an indicator for whether responses could differ depending on views about the efficiency of the relevant markets.[8]

We find that the probabilities that respondents are uncertain or reject the consensus view plummet when the academic literature is large. Compared to having no past literature, when

[8] The data assignments were simply a judgment call, and are listed in the Online Appendix.

there is a large literature the fraction uncertain is 25 percentage points lower (compared to a mean of 25 percent), while the fraction rejecting the consensus is 8 percentage points lower (compared to a mean of 6 percent). Also, when responses could vary depending on the degree of confidence in efficient markets, the chance of rejecting the consensus is a little over 2 percentage points higher. There is also less uncertainty when the statement implicitly raises efficiency or distributional issues.

Note that the coefficients on the individual characteristics are robust to the inclusion of the additional controls in the second specification. In principle, individual non-response patterns could have been correlated with characteristics of the questions. By finding no change in these individual coefficients, we infer that non-responses are largely unrelated to characteristics of each question.

As a more complete control for differences across questions, the third column includes separate dummy variables for each question. Most of the coefficients on the individual characteristics remain qualitatively unaffected. The exception is that experts are now found to be significantly less likely to be uncertain.[9] Of course, experts know the existing literature much better, so that this effect is entirely consistent with our null hypothesis that opinions simply reflect the state of the existing academic literature.

The only statistically significant deviation from homogeneous views, therefore, is more caution among women in expressing an opinion. Personality differences rather than different readings of the existing evidence could well explain these gender effects.

We certainly see differences among respondents in the confidence they assign to their responses. To document these differences, we regress an individual's reported confidence on each question against a set of individual characteristics. We include the same individual controls

[9] This final specification also includes an additional control variable, "average confidence", which we discuss later.

used above, along with dummy variables for the university where the individual is currently employed and also dummy variables for the university where they received their Ph.D. In addition, we added the indicators for questions dealing with market efficiency or distributional issues, and controls for the size of the literature on the question, expecting respondents to be more confident when the literature is large.

Table 2. OLS Estimates of the Determinants of Confidence.

Variable	Coefficient	t-test
Female	-.33	-2.47
Washington	.49	3.67
Vintage: older	-.35	-2.04
mid-career	.84	6.87
Expert	.68	6.59
From: Chicago	1.51	2.90
Harvard	.81	1.73
MIT	1.31	2.74
Princeton	-.21	-0.40
Stanford	.28	0.57
Yale	-.26	-0.51
At: Chicago	-.26	-1.37
Harvard	-.28	-1.42
MIT	.55	2.68
Princeton	.74	3.42
Stanford	.49	2.47
Yale	1.04	5.02
Literature: int. micro	1.16	8.40
large	1.14	8.41
medium	0.24	1.92
Perfect markets?	-0.07	-0.53
Redistribution?	-0.01	-0.08
R-squared	.114	

Notes: 2,612 observations. Confidence ranges from 1 to 10, with a mean of 6.1 and a standard deviation of 1.0. The omitted category is a young male, without experience working in Washington, without any expertise on the topic, with a Ph.D. from Berkeley, now teaching at Berkeley, facing a question with no past literature.

Results appear in Table 2. Respondents are dramatically more confident when the academic literature on the topic is large. Not surprisingly, experts on a subject are much more confident about their answers. The middle-aged cohort (the one closest to the current literature) is the most confident, while the oldest (and wisest) cohort is the least confident.   Those who have worked in Washington do show some tendency to be more confident, while women have a tendency to be less confident. Respondents who got their degrees at Chicago are far more confident than the other respondents, with almost as strong an effect for respondents with Ph.D.'s from MIT and to a lesser extent from Harvard. Respondents now employed at Yale and to a lesser degree Princeton, MIT, and Stanford seem to be more confident.

Given such clear differences in the confidence expressed by members of the Panel, we therefore included a confidence measure in the last column in Table 1, defined to equal the individual's average confidence expressed on all questions *other than* the current question, to avoid possible reverse causation. The intent here is to control for personality differences, so that the remaining variation is more likely to be due to different interpretations of the academic literature. As expected, individuals with a higher mean confidence are less likely to answer that they are uncertain. More confident individuals are also more willing to disagree with the consensus. However, the inclusion of this variable does little to change the coefficients on gender, leaving open the possibility that women tend to interpret the existing literature differently.

The above results estimate the degree to which individuals differ in their likelihood of joining the consensus view on a topic. They provide no evidence, though, on whether individuals group into different schools of thought, agreeing on when to disagree with the consensus. As a first test for the existence of different schools of thought, we examine whether people with similar characteristics provide correlated responses, controlling for the expected pattern of responses coming out of Table 1.

In implementing this test, we created a one-dimensional measure of responses with "1" indicating the consensus response, "-1" indicating the opposite response, and "0" indicating a response of "uncertain". The expected value of each individual's response is based on the above multinomial logit model, including just the question dummies.

We then regress the constructed correlation between the responses for each pair of individuals against a vector of dummies capturing shared characteristics: shared experience working in Washington for the same political party, shared experience working in Washington but for different political parties, being faculty in the same Department, having Ph.D.'s from the same school, being members of the same cohort, having the same gender, and sharing expertise.

Table 3. OLS Estimates of the Determinants of Correlation in Responses.

Variable	Coefficient	t-test
Both Washington / same party	.068	1.84
Both Washington / diff. party	.020	0.52
Same expertise	.038	2.30
Both older	-.057	-1.87
Both mid-career	.019	0.94
Both younger	-.028	-1.69
Same university	.008	0.45
Same Ph.D. program	.008	0.57
Same gender	-.018	-1.37
R-squared	0.021	

Notes: 820 observations. The average correlation over all respondent pairs is 0.083. The omitted category is pairs of individuals lacking any of these common traits, e.g. from different cohorts, with different genders and different fields of expertise.

Results are reported in Table 3. Here we find that pairs of individuals sharing experience working in Washington for the same political party tend to agree with each other. However, if two individuals both worked in Washington but for different political parties, they are no more likely to disagree than are any two members of the sample, undermining any support for polarized views even among those who have worked in Washington. The agreement among those working for the same party could simply be shared knowledge about particular policy issues.

Among the other findings, those sharing expertise are much more likely to agree with each other. More striking, though, is that two younger individuals are more likely to disagree than are two individuals from different cohorts; the same is true for two older individuals. Plausibly, some of the less risk-averse young respond to the professional payoff to successfully challenging existing views. When one is old, there may simply no longer be any professional cost from disagreeing with the consensus. Non-results are also intriguing: there is no tendency for those from the same Ph.D. program or those currently teaching at the same school to agree with each other, contrary to common presumptions.

The results so far do nothing to identify panel members with conservative versus liberal views. To make progress here, we tried two different approaches. In the first, we looked for clusters in the data, using a variety of distance-based clustering methods and several specifications for each method.[10] The data show no systematic evidence of clustering into two or even a few roughly equal-sized camps, yielding only weak evidence that three or four idiosyncratic individuals differ from the rest of the sample, and from each other.

As a second approach, we focused on the subset of questions that we flagged previously as likely to generate disagreements due to distributional concerns or due to debates on the possible importance of market failures. Here, we went a step further and classified one answer as consistent with a "Chicago price theory" perspective, and the other answer as likely reflecting distributional concerns or concerns about market failures or behavioral anomalies.[11] Our null hypothesis was that individuals are homogeneous and therefore act based on the response rates

[10] We explored both hierarchical clustering and a procedure based on random-partition distributions. We would like to thank David B. Dahl for advice and help in the use of clustering techniques.

[11] These classifications can be found in the Online Appendix.

for the Panel as a whole, sometimes giving the "Chicago price theory" response and sometimes disagreeing with this response or expressing uncertainty. If there were liberal and conservative camps, in contrast, some would consistently give the "Chicago price theory" response and others would consistently give the opposite response. Here, the data were entirely consistent with the null hypothesis: implementing a chi-squared test of homogeneity (see footnote 4) yields a p-value of 0.70. There is no support for different camps in our data, only idiosyncratic views among some Panel members.

II. Summary

The immediate finding when examining the responses of this Panel of economists is the remarkably high degree of consensus. The few disagreements that exist seem to arise largely when the academic literature on a question is small or non-existent, allowing differing prior beliefs to remain an important determinant of posterior views. Economists in mid-career seem to be closer to the professional consensus than either younger or older economists. Women tend to be more cautious in taking a stance, while those with experience working in Washington have fewer such inhibitions. Surprisingly, there are no detectable systematic differences in views across Departments, or across school of Ph.D. In addition, there is no evidence to support a conservative versus liberal divide among these Panel members, at least on the types of questions included so far in the surveys.